



News on Research, Products and Solutions for Learning and Education

innovations



inside:

**Automated
Scoring
Technology**

Data Explorer

CriterionSM
Online Writing
Evaluation



Listening. Learning. Leading.



Editorial Staff

2 Project Director, William Petzinger
Executive Editor, Rich Pliskin
Contributing Writers, Janet Levy, Amanda McBride,
 Martha Mendez
Senior Designer, Marita Gray
Production Manager, Ellie Patounas

Editorial Board

Richard Coley, Director, Center for Policy
 Evaluation & Research
Laura Delate-Miller, Director, Creative Services
Gregg Driben, Associate Vice President, Elementary &
 Secondary Education
Anne Gale, Vice President and Corporate Secretary
Francis D. Gómez, Executive Director,
 Corporate Communications
Daniel Johnsen, Assessment Specialist,
 Language Skills Assessments
Anne-Marie Matyi, Marketing Specialist
William Monaghan, Manager, Research Communications
Laurie Russell, Director, Community Services
Tauni Sauvage, Director, International Marketing
Linda Scatton, Senior Program Administrator
Edward Shea, Director, Research
Lew Shumaker, Senior Diversity Consultant
Kelly Smedley, Director, Client Services &
 Project Management
Barbara Vilkomerson, Managing
 Director, Higher Education
Daniel Wakeman, Chief Information
 Officer, Elementary & Secondary Education
Eric Waxman, Director, Consulting
 Services, Strategic Workforce Solutions
Peter Yeager, Manager, External Relations



A Letter From Kurt Landgraf

Information is power.

We believe in education because information and knowledge have the power to improve people's lives. But information is also a powerful tool in education itself. Educational assessment, after all, is about generating information on student learning: The more teachers know about who's learning and who's not, the more able they are to improve student performance.

One of the most powerful ways to put information to educational use is through information technology, which has proved a fantastic boon to classroom practice and public policy. In this issue of *ETS Innovations*, we examine some of the ways in which information technology is reshaping teaching and learning throughout the world.

Our automated scoring technologies, for instance, are making it possible for students to get their test scores in a fraction of the time required by manual scoring.

Scoring technologies are also enhancing the learning process itself. One example is our *Criterion*SM Online Writing Evaluation service, which gives students immediate feedback on their writing skills. Another is our *SpeechRater*SM service, through which students who use our TOEFL® Practice Online service can get instant feedback on their speaking skills to help them prepare for the TOEFL test.

We've even developed a test, called the *Information and Communication Technology Literacy Assessment*, that uses innovative technology to measure students' proficiencies with technology in practical and academic settings.

We're pleased to share with you information on these and other innovative educational technologies that ETS has designed and developed. Information is power. And information technology is helping to put the power of education into more hands than ever.

Regards,

Kurt M. Landgraf
 President and CEO

inside

- 3** What's Inside: New Technologies Improve Automated Scoring
- 6** Learning to Write, Writing to Learn
- 8** The TOEIC® Test Now Measures 4 Communication Skills
- 9** The TOEFL® Test in China: Yesterday, Today & Tomorrow
- 10** The Data Explorer: Making Sense of Statistics, Improving Student Achievement
- 12** When Mastering MP3 Files Isn't Enough: Helping Students Become Technologically Literate
- 15** ETS Classifieds

Let us know what you think about *ETS Innovations*, including ideas and issues you'd like to see explored in these pages. Send our editors a note at

InnovationsMailbox@ets.org

In the meantime, please visit the ETS Web site at
www.ets.org.

Copyright © 2006 by Educational Testing Service. All rights reserved. ETS, the ETS logo, TOEIC, TOEFL, GRE, the GRADUATE RECORD EXAMINATIONS and E-RATER are registered trademarks of Educational Testing Service (ETS) in the United States of America and other countries throughout the world. C-RATER, SPEECHRATER, M-RATER, CRITERION and SCOREITNOW! are trademarks of ETS. 2249



What's Inside

New Technologies Improve Automated Scoring

A student in Paris speaks into a computer microphone and in minutes gets a score on her English-speaking proficiency as part of a new service called [TOEFL® Practice Online](#). Like students using the service in Tokyo, Dubai and countless other locations around the world, she'll be able to use this experience to better prepare for the [TOEFL® iBT](#) Speaking test.

What makes this immediate feedback possible is the *SpeechRaterSM* scoring engine designed by ETS research scientists. It's the newest entry in a suite of automated scoring technologies in development at ETS that are creating innovative solutions for teaching and learning.

And TOEFL iBT, the Internet-based version of the [TOEFL test](#), is just one place where new technologies are making a big difference.

TOEFL Practice Online also uses another of ETS's automated scoring technologies — the *e-rater®* scoring engine. The *e-rater* engine provides students immediate scoring on their practice essays — in this case, to help them prepare for the expanded [TOEFL iBT](#) writing section.

Driving innovation

Automated scoring technologies are also being used to score operational assessments, that is, nonpractice tests for which an official score is given. "In addition to *e-rater*, schools are also using the *c-rater™* engine, which scores short responses for content, and the *m-raterSM* engine, which scores mathematical expressions and graphs," says Kim Prickett, Senior Project Manager in ETS's Elementary & Secondary Education Division.

"These new technologies reduce score turnaround times from weeks to just days,

enabling educators to more quickly use the test data," Prickett says.

The *e-rater*, *c-rater*, *m-rater* and other next-generation technologies are central to ETS's mission to advance quality and equity in education.

"An important part of ETS's mission is to drive innovation — including technological innovation — that continually advances assessment, learning and teaching," says Ida Lawrence, Senior Vice President of Research & Development at ETS. "The automated scoring technologies are excellent examples of practical, research-based tools that help students and teachers."

Using speech recognition and processing technologies, the *SpeechRater* engine scores students' spoken responses to test questions by combining the evaluation of several important language features

— pronunciation, fluency, vocabulary and grammar. Together, these features cover part of the scoring criteria used by human raters to score answers on the actual TOEFL iBT Speaking test.

“Even though the *SpeechRater* service currently focuses only on a core set of language-use features and the program’s ability to evaluate topic development is in its infancy, we see it as an evolutionary capability to meet a real business need in any country where the need arises,” says Gerard Mannarino, ETS Program Director for TOEFL Practice Online.

“The *SpeechRater* service is an emerging capability, so we’re starting out in a practice environment to help us mold and refine the technology and match it with users’ needs,” adds R&D Research Director David Williamson, who leads the development of ETS’s automated scoring capabilities.

“While the technology helps non-native speakers of English more effectively prepare for the TOEFL iBT, it will also provide data that will allow ETS scientists to make improvements to the capability,” Williamson says.

Shaping the future

ETS’s research agenda and the reaction of users to this initial version of the *SpeechRater* engine will help shape development of later versions that will more fully address speech delivery.

“While Version 1 focuses on providing summary scores, in subsequent versions we’ll combine delivery, language use and content for task-based feedback to help students evaluate their own performance,” Williamson says.

“Our goal is a version in which speech scoring replaces the role of human evaluators for assessments in which only group-level scores are provided, or replaces one of two human evaluators for assessments for which an individual receives a score.”

ETS Research Scientist Xiaoming Xi notes that the features of Version 1 were based on input from ETS test-content specialists and



▲ ETS Research Scientist Xiaoming Xi says key features of Version 1 drew on input from language-learning experts.

language-learning experts and practitioners in the United States and abroad.

“The scoring model has been endorsed by both content specialists and statistical experts,” she says. “We continue to carry out rigorous content and technical reviews, and we are communicating to potential users and the general public what this first version can and cannot do.”

Scoring writing quality

While development of the *SpeechRater* service is in the early stages, the *e-rater* technology is a more fully developed computer program that automatically scores essays primarily on the basis of writing quality. It uses natural language processing technology to evaluate the characteristics of an essay, including grammar, usage, mechanics and development.

The *e-rater* service is being used in several assessment and learning programs. ETS’s award-winning *Criterion*SM Online Writing Evaluation service (see story on page 6), for instance, uses the *e-rater* technology to evaluate students’ skill levels, allowing instructors to benchmark writing, adjust instruction and track progress.

Hundreds of thousands of students, from grade 4 through adult, use the *Criterion* service in school districts, colleges and universities throughout the world.

The *ScoreItNow*TM writing service for the GRE[®] test also uses the *e-rater* engine to score practice essays for the writing section of the Graduate Record Examinations[®] General Test. And ETS is developing several other writing practice programs that will also use the *e-rater* service.

ETS devotes considerable attention to the structure of the test items that the *e-rater* engine will score, as well as to the scoring guidelines for them, Williamson says. ETS administers items in a pretest to a minimum of 500 students. The items are then scored by specially trained professionals, preferably twice.

This process forms the basis for building *e-rater* automated scoring models, for which the results are cross-validated against an additional 500 student essays.

“Because the *e-rater* engine does not score for content or ‘read’ the essays, we have established standards for item and rubric development, as well as for the empirical relationship between human and automated scores,” Williamson says.

Scoring short responses for content

A newer capability is the *c-rater* engine, which automatically scores short, fact-based, English-language responses that are typically between two and five sentences long.

The *c-rater* engine measures respondents’ understanding of specific content by using automated natural language processing techniques to determine whether responses contain evidence that the test taker has reproduced a specified concept in an answer.

“The *c-rater* engine has a feature that matches examinees’ short responses to concepts expressed in a linguistic model,” says ETS

The scoring model has been endorsed by content specialists and statistical experts.

Xiaoming Xi

Associate Development Scientist Na-Rae Han. "Each item has its own specific scoring guide that spells out how credit may be earned for the item and what constitutes a particular score.

"ETS test developers build the linguistic model based on a scoring rubric and a sample of human-scored responses, and they refine the model until it is sufficiently accurate based on the pretest sample of human-scored responses," she says.

"Like ETS's other automated scoring technologies, we are continually improving *c-rater* as our scientists learn more about automated scoring of constructed response tasks."

Scoring constructed response math items

In testing programs, the *c-rater* service is often used in conjunction with another developing technology, the *m-rater*SM engine. According to ETS Business Systems Lead Analyst Jim Fife, the *m-rater* engine automatically scores examinees' computerized constructed responses that can be in the form of equations, graphs or numbers.

The *m-rater* technology can be used with ETS's Equation Editor and Graph Editor, which allow a test taker to enter an equation or to graph points, lines, broken line segments, and curves. For equations, the *m-rater* engine determines the mathematical equivalence of the examinee's response to the correct response as given in the scoring key.

As with all automated scoring, there must be substantial coordination between item development and scoring. "It's important that the item be constructed in a way that the *m-rater* engine can score it," Fife says. "For example, if the item asks for two equations, then there need to be two separate answer boxes, one for each equation."

Strengths and limitations

Automated scoring has strengths that make it very attractive, says Williamson. "Score turnaround is high on every client's list of

requirements, and automated scoring allows for a dramatic reduction in turnaround times. *E-rater* scores take seconds to produce; the *SpeechRater* engine is expected to reduce the scoring lag from five days to just minutes.

"And while a human scorer uses a rubric to produce a summary score, it's not transparent how that score was generated," Williamson adds. "Automated scoring provides a detailed account of the score rationale that can be compared with expert opinion about the relative importance of scoring features."



▲ ETS Research Director David Williamson says automated scoring provides a detailed account of the score rationale.

"The scoring model is then uniformly and accurately applied in scoring all the responses."

But even under ideal conditions, automated scoring can fail. "You must always recognize that there are limitations, including the scoring-engine development, item and rubric design and, most importantly, the quality of human scoring practice that the automated scoring depends on," Williamson says.

"To mitigate these factors, we provide for additional review of item and rubric development, pretest more items than required, and ensure high standards for human scoring as a precursor to automated scoring."

Score
turnaround is
high on every
client's list of
requirements.

David Williamson

Preparing for the future

2006 has been a dynamic year for ETS's automated scoring activities. "ETS R&D made a major investment to ensure the quality of our capability development and the evaluation of the operational use of these technologies," Williamson says.

To advance the field of automated scoring, Williamson recently edited *Automated Scoring of Complex Tasks in Computer-Based Testing* with ETS Principal Research Scientist Isaac Bejar and Robert Mislevy of the University of Maryland. Another topical ETS-related book is *Automated Essay Scoring: A Cross-Disciplinary Perspective*, which ETS Principal Development Scientist Jill Burstein edited with Mark Shermis of Florida International University. Lawrence Erlbaum Associates published both books.

Williamson anticipates ongoing and substantial advances in ETS's technical capabilities around automated scoring.

"With researchers now using operational data to inform development, the question is not whether we can improve upon our automated scoring capabilities, but rather at what pace can we maintain incremental improvements in the construct representation and performance of automated scoring," he says.

"As we continue to develop the next generation of scoring capabilities, we expect to come up with new applications and to re-evaluate what we previously thought could only be scored by human graders." 🌟



Learning to Write, Writing to Learn

**ETS's *Criterion*SM Online
Writing Evaluation Service
Comes to Writing
Labs Around the World**

The term “automated scoring” once conjured images of machines scanning reams of multiple-choice answer sheets in search of little ovals filled in with pencil marks. But grading writing assignments? Only a few years ago, that would have sounded like science fiction.

With its *Criterion*SM Online Writing Evaluation service, ETS has transformed science fiction into classroom reality.

Since 2002, the *Criterion* service has been improving the academic experiences and professional lives of students and teachers around the world, from elementary and secondary schools to colleges and universities. Along the way, it's won industry plaudits as an innovative learning solution.

This past May, the *Criterion* service won the Software & Information Industry Association's coveted Codie Award for Best Instructional Solution for English-Language Acquisition. It was also a finalist in the category of Best Instructional Solution for Language Arts/English — Secondary.

How it works

The *Criterion* service is a learning tool that uses automated scoring technologies to

evaluate students' essays online. Despite the sophistication of the underlying technology, ETS designed the software to be easy for students and teachers to use.

At the heart of the service lies the *e-rater* technology. Developed by ETS researchers, the *e-rater* scoring engine enables the *Criterion* service to provide students with a holistic score within seconds of submitting an essay online. Students also receive annotated feedback in real time on the major “trait categories” of grammar, usage, mechanics, style and organization and development as well as 36 specific subcategories.

After teachers assign a writing task, students can access the *Criterion* service from any location with an Internet connection and type in their work. Students can also use the service to revise previous work with help from any of six online writer's handbooks, including a bilingual Spanish/English guide, that provide writing advice and examples.

Regardless of whether the work is new or revised, within 20 seconds of submitting a finished draft, students receive their holistic score and detailed feedback.

'Incredibly helpful feedback'

"The feedback and holistic scores are incredibly helpful to students, especially since students get them in a low-stress, learning environment," says Linda Reitzel, Director of Development for the *Criterion* program in ETS's Elementary & Secondary Education Division.

"And because the software saves students' first and most recent essay submissions with their score and feedback, students tend to treat the writing experience as a personal challenge to improve."

Since it launched the *Criterion* program in 2002, ETS has expanded the service's library of essay topics. More recent versions that are in use, including the current Version 6.2, can also score essay topics that instructors create themselves, customized to teachers' curriculums.

Says Reitzel, "One recent note we received came from a high school teacher in Indiana: 'With the *Criterion* technology, my gifted students are challenged, my average students are engaged, and my weak students are specifically identified and supported.'"

Yet even the most sophisticated technology can have its limits. With the *Criterion* service, the *e-rater* scoring engine doesn't read essays for context and content

correctness. So, for example, the software would view "Man bites dog" as grammatically acceptable as "Dog bites man."

Still, teachers have been enthusiastic about the service, which they view as a tool that helps them identify students' trouble spots and allows them to devote more time to improving content and essay development versus marking errors.

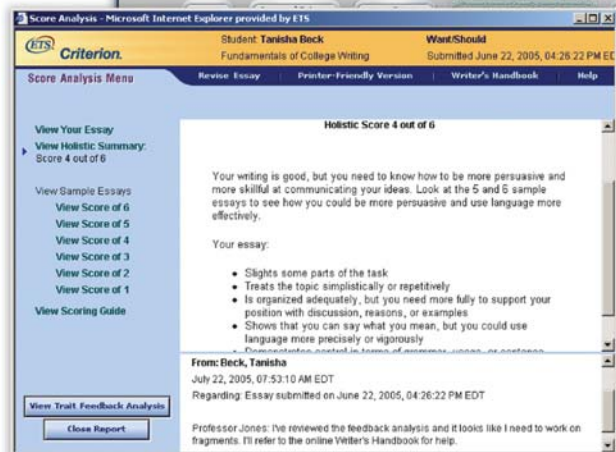
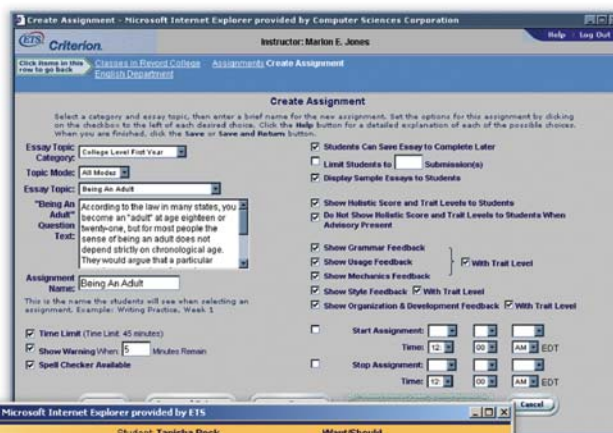
Teachers also say that the service has enhanced their one-on-one interactions with students. For example, the software allows teachers and students to post notes on a message board for one another, and teachers can insert pop-up notes in students' essays.

Teachers as advocates, not adversaries

Because the service, not the teacher, points out errors and trouble spots, students tend to view the computer as the critic and their teacher as a resource for improving. "Teachers tell us that they are now seen as the advocates, not the adversaries," Reitzel says. "And they enjoy that!"

In addition to its success in the United States, the *Criterion* service has been embraced by elementary, secondary, post-secondary and language schools around the world, with customers in Australia, Canada, Japan, Nepal, Singapore and the United Arab Emirates.

In the United States, more than 1,000 districts, 3,000 schools, and 9,500 instructors use the service, with more than 1 million essays submitted over the past year. As of June 2006, 42 U.S. states were using it for elementary and secondary education, and 43 for higher education.



More essay topics

Bolstering the service's use in higher education are the 90 essay topics ETS added in August, which more than doubles the number that instructors can use. Forty of the new topics present business scenarios. A full library of persuasive, expository, descriptive and narrative essay topics is also available. Faculty can also create their own topics.

"As with the elementary and secondary education market, an immeasurable benefit of the *Criterion* service is the power and flexibility it offers educators," says Linda Tyler, Executive Director of the Postsecondary Solutions Area for Higher Education at ETS.

"Rather than boxing them in, the service offers them open-ended opportunities to target instruction and improve their students' writing."

For more information about the *Criterion* Online Writing Evaluation, go to www.ets.org/Criterion. Be sure to view either the K-12 *Criterion* Tour or the Higher Education demo. ✨

Students treat the writing experience as a personal challenge to improve.

Linda Reitzel

The TOEIC® Test Now Measures 4 Communication Skills

Listening is not enough.

Reading is not enough.

Speaking is not enough.

Writing is not enough.

Today's highly competitive and increasingly global workplace requires a workforce competent in the full range of English-language communication skills. That means being able to hold complex conversations, conduct meetings, compose effective business correspondence, and read technical reports.

In other words, more than ever, employees competing for jobs need to be able to listen and speak as well as read and write in English — whether or not English is their native language.

As a result, businesses competing in the international arena have been clamoring for tools to help them better assess current and prospective employees' English-language skills.

Responding to these demands, ETS has enhanced and expanded the [TOEIC® test](#). The new TOEIC suite of tests includes authentic reading and listening tasks as well as much-anticipated speaking and writing tests.

ETS will be using its latest assessment technologies, including Internet-based testing and the [Online Scoring Network](#), for delivering and scoring the test. The move enables ETS to assess the skills businesses want and, at the same time, enhance test-takers' experience with the test.

"Although the TOEIC test is already considered to have set the global standard for assessing communicative workplace English ability, we are keenly aware of the corporate world's changing — and growing — needs to gauge employees' communication skills," says Paul A. Ramsey, Senior Vice President of ETS's Global Division.

"We are determined to meet those needs as we always have: with the most fair, valid assessments possible."

That's exactly what ETS is doing with regard to the TOEIC testing program.

It's about the real world

In May 2006, ETS launched the new TOEIC Listening and Reading test in Japan and Korea. The new test is a redesigned version of the TOEIC Listening and Reading exam taken by 4.5 million people each year. The new test includes more authentic English-language reading and listening tasks to better reflect the English used in the global workplace.

Along with revamping the exam, ETS has revised the score report for the TOEIC test. The new report provides test takers with feedback to help them better assess their strengths and weaknesses, which helps guide them in their learning.

In December 2006, ETS will begin administering the TOEIC Speaking and Writing tests for the first time. These assessments will measure test-takers' ability to use spoken and written English.

On the Speaking test, candidates will be required to read a text, describe a picture, respond to short questions, respond to a complaint, and express an opinion. They'll be evaluated on their pronunciation, intonation, grammar, vocabulary and the relevance and completeness of their responses.

With the Writing test, test takers are asked to write a sentence based on a picture, respond to a written request, and write an opinion essay. They will be evaluated on overall organization, appropriate and precise use of grammar, and vocabulary. Both tests will measure a wide range of skills.

In advance of launching these new tests, ETS has introduced the [TOEIC Speaking and](#)



In May 2006, ETS launched the new TOEIC® Listening and Reading test in Japan and Korea.

Writing Official Practice Tests. The practice tests, which are identical to the operational tests, give examinees the opportunity to practice online for the TOEIC Speaking and Writing operational tests in a simulated testing environment. Test takers can practice anytime, anywhere — at home, at a language lab or at work — via the Internet. It was another technological advance — ETS's Internet-based test delivery system, or iBT — that made it possible for ETS to develop the new TOEIC Speaking and Writing tests.

iBT enables ETS to simultaneously deliver tests over the Internet to computers located anywhere in the world. Perhaps more importantly for the TOEIC Speaking test, iBT makes it possible for ETS to capture speech samples from test takers.

"Before iBT, the only way we could capture test-takers' speech was with tape recorders,

which was a costly and very cumbersome process," says Feng Yu, Director of ETS's Global Product Management Group. "Now, we can offer an accurate speaking test and do so efficiently, affordably and in a fair manner."

In addition, ETS-trained and certified raters will score the responses via the ETS-developed Online Scoring Network (OSN). OSN is a fast, secure computer-based system that allows qualified raters to evaluate responses online, either from home via the Web or at an OSN center.

Catering to customer needs

The technologies ETS is employing to deliver and score the new TOEIC test have also permitted ETS to be more responsive to customers' needs. "We want to minimize the inconvenience of test taking," Yu says. "To do that, we have to have technology that



▲ The new technology minimizes the inconvenience of test taking, says Feng Yu, Director of ETS's Global Product Management Group.

is flexible enough to accommodate local cultures and needs."

"The new TOEIC test, along with its new features, accomplishes this — all at a competitive cost and with ETS's high quality standards." 🌟



The following are excerpts from a speech by Paul A. Ramsey (above center), Senior VP of ETS's Global Division, at the International Conference for Testing Methodologies and Theories in Beijing on Sept. 14, 2006.

In just a few months, [ETS] will mark the 25th anniversary of the first administration of the TOEFL® test in China. ETS has worked with the Ministry of Education and education leaders in China for 25 years, helping to open the doors to academic institutions for Chinese

students, and to promote and support English-language learning.

The Chinese students who have taken the TOEFL test and studied in other countries gained the benefit of an international education. The countries that have hosted them learned from them also. They learned more about your culture and way of life. They learned about the unique talents, ideas and perspectives that Chinese students bring with them when they study abroad.

Tomorrow, ETS will mark another significant milestone in our quarter-century relationship with China and its students in that we will administer the first Internet-based TOEFL tests in this country.

How English is actually used

Test delivery systems are changing and those changes have significant implications for test takers, education leaders,

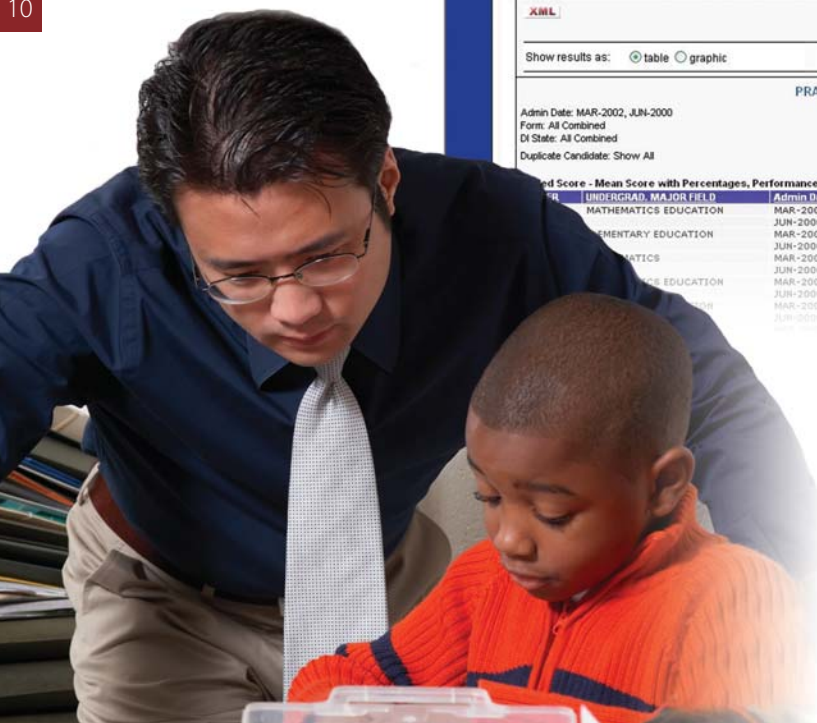
teachers and score users in China and other countries. In the case of the English language, the Internet enables us to measure as never before the way English is actually used on campuses and in real life.

In college classrooms, there are no writing or listening test modules, but there are lectures, assignments and reading requirements that require an integrated response. In business, there are no reading sections, but there is the English language just as there is the Chinese language. These are living, comprehensive languages people use to communicate, to conduct international commerce.

Because the new test raises the bar on English-language skill measurement, it will also change significantly the way that English is taught to non-native speakers. That will benefit every student — those hoping to study abroad, those seeking to enhance their English skills at home, and those wanting to advance their careers.

(The TOEFL test in China continued on page 15)

The TOEFL® Test in China: Yesterday, Today & Tomorrow



The Data Explorer:

Making Sense of Statistics, Improving Student Achievement

When state education officials in the United States want to learn how their students are doing in math compared with students in other states and around the nation, there's an easy way for them to find out.

Using the ETS-developed Data Explorer, they can access a wealth of information and create highly customized reports with vital data on academic achievement, trends and group differences that they can share with parents, teachers, policymakers and the media.

ETS developed the Data Explorer to help the U.S. Department of Education make the 900 million summary statistics from the National Assessment of Educational Progress (NAEP) — commonly called the Nation's Report Card — accessible to the public.

Available on the NAEP Web site, <http://nces.ed.gov/nationsreportcard/nde/>, the [NAEP Data Explorer](#) lets users explore the results of decades of assessment of students' academic performance and information about factors that may be related to their learning.

"The NAEP Data Explorer has proven to be a valuable tool for anyone who has questions

about what the nation's students know and can do," says John Barone, who heads ETS's Center for Data Analysis Research.

Customer customization

"The database delivers custom tables of results from all major national and state assessments conducted for NAEP since 1990, including results from 6 million students assessed in eight subjects and three grades," Barone says. "It presents data summarized on more than 1,000 variables containing background data from students, teachers and schools.

"These are highly complex data sources," he says. "But directed by a few clicks of the computer mouse, the Data Explorer can dig down into the data to present information in a clear and understandable form."

ETS has also produced customized versions of the tool for such clients as Statistics Canada, Canada's central statistical agency, and the U.S. state of New Jersey. Statistics Canada uses the tool to analyze and report data from the [International Adult Literacy Survey](#), while New Jersey applies it to the [New Jersey Assessment of Skills and Knowledge](#)

(NJ ASK) administered to third- and fourth-grade students.

"One of the Data Explorer's most valuable features is this adaptability," Barone says. "It can help diverse clients in other countries make sense of many different sets of statistics and variables."

Analyzing NAEP data

According to ETS Senior Program Administrator Debbie Kline, the NAEP Data Explorer lets users:

- report data aggregated by student groups, states and school districts
- summarize achievement data across years
- show trends and group differences
- compare results with the nation, the region and other states
- show results in the form of tables, bar charts, trend graphs and maps
- drag and drop tables into Microsoft applications such as Word, PowerPoint and Excel

"The tool enables users to disaggregate and examine test data to pinpoint academic performance of subgroups and address differences and shortcomings," Kline says.

"State administrators tell us that's crucial in this era of educational accountability. They're also using the Data Explorer to provide policymakers with reports that give a clear picture of student achievement to highlight essential funding needs."

The ability to address customers' specific needs is a boon, Kline says. "The NAEP Data Explorer provides information on states and selected school districts, but the version that ETS customized for New Jersey allows users to view any combination of performance measures that are available for a number of group variables, including gender, ethnicity, school district socio-economic status, and school performance levels," Kline says.

Just ASK New Jersey

In 2003, ETS developed a version of the Data Explorer to help show how well students taking the *New Jersey Assessment of Skills and Knowledge* are meeting the requirements of the No Child Left Behind Act. NCLB is the U.S. law that aims to toughen standards and accountability in public education and close academic achievement gaps among students of different racial, ethnic and economic backgrounds.

The NJ ASK exam tests language arts literacy, math and science. Kline notes that the NJ ASK Data Explorer now contains two years of data, allowing comparisons between years. "As more data are added, the tool will become even more useful in helping to inform state educational progress."

A version of the tool is being used to analyze national and

international data as well. ETS developed a version of the Data Explorer for Statistics Canada to help users analyze and report data provided by the International Adult Literacy Survey (IALS). The database is available in English and in French on the Statistics Canada Web site, www.statcan.ca.

The ETS Literacy Data Tool, at www.ets.org/etsliteracy, offers the same features and allows users to look up data from the *Adult Literacy and Lifeskills Survey (ALL)* and the IALS according to a variety of categories, such as education level, demographics, language background, and labor force experiences. For countries that participated in both IALS and ALL, information is available to allow comparisons over time.

Learning about learning

A joint program of Statistics Canada and the U.S. Department of Education's National Center for Education Statistics, IALS also provided the world's only source of comparative data on adult skills and the relationships of these skills to social, educational and labor market outcomes.

The ALL survey maintains the measures of literacy used in IALS but adds two new domains: numeracy and problem-solving, or analytical reasoning.

The tool is helping researchers, institutions and policymakers access and use information from these surveys to highlight literacy trends and identify learning needs around the world.

Early in 2007, the Literacy Data Tool will allow users to access data from the Adult Education Program Study that will be hosted on the ETS Web site.

ETS has also produced customized versions of the data tool for such clients as Statistics Canada.

Funded by the U.S. Department of Education's Office of Vocational and Adult Education with the Center for Education Statistics and the Department of Labor, the 2003 study provided data on a nationally representative sample of adult education programs and the services they offer.


What's next for the Data Tool

In the study's second phase, the Adult Education Program Study's Learner Survey assessed the literacy skills of a nationally representative sample of adult learners who participate in these programs.

"The measures that were used allow us to compare these adult learners with the household population participating in the Adult Literacy and Lifeskills Survey," says Irwin Kirsch, Director of ETS's [Center for Global Assessment](http://www.ets.org/globalassessment).

"Moreover," Kirsch says, "Hispanic adults participating in these programs were tested in either English or Spanish, allowing us to compare literacy performance in both their native and non-native languages. This is the first time a nationally representative sample of adult learners has been studied and compared with the household population."

Noting the multiple uses of the tool to date, Barone says the Data Explorer's adaptability is one of its greatest assets. "We've been able to build customized versions of the tool for clients who are using it to mine important information from a variety of large databases," he says.

"As the demand for good data analyses and effective ways to present them continues to grow, we expect to develop new versions of the Data Explorer to help governments, states, foundations and business clients make more effective use of their statistics." 



"We expect to develop new versions of the Data Explorer to help governments, states, foundations and business clients make more effective use of their statistics."

John Barone, ETS Center for Data Analysis



iSkills™

Listening. Learning. Leading.

Educational Testing Service
Rosedale Road
Princeton, NJ 08541

Announcement of product name change:

The *ICT Literacy Assessment* is now called the *iSkills™* assessment.
All references to the ICT Literacy Assessment in the following document
apply to the *iSkills* assessment.

◀ *Stephanie Brasley,
Information Literacy
Coordinator at UCLA,
works with student Amber
Ramirez on the ICT
Literacy Assessment.*

When Mastering MP3 Files Isn't Enough:

Helping Students Become Technologically Literate

Today's college students can navigate Web sites faster than you can say "podcast," and consider sending a text message as routine as making a phone call. So why does research characterize these students as less than technologically savvy?

The answer lies in something called ICT, or information and communication technology, literacy, which involves the ability to combine cognitive and analytical skills in

the context of technology. Despite their ease with technology, college students generally exhibit low levels of ICT literacy.

"Feedback from librarians and faculty members has shown that while most college-age students can use technology, they don't necessarily know what to do with the content the technology provides," says ETS Senior Research Scientist Irvin Katz.

As a result, Katz says, students are able to create PowerPoint slides loaded with animation and hotlinks, but lacking persuasive impact — resulting in a final product that may look good but contains little substance.

"Young people have generally used technology for entertainment and purchasing," adds Teresa Egan, ETS's ICT Literacy Project Manager. "But technological proficiency involves so much more. To achieve ICT

literacy, a person must be able to research a topic well, identify credible data, and then prepare a compelling, cohesive argument based on the findings."

How real is the problem? Consider the following ETS data on college students' technological proficiency:

- When asked to evaluate a set of Web sites, only 52 percent judged the objectivity of the sites correctly.
- When selecting a research statement for a class assignment, 48 percent chose one that was too broad, albeit reasonable, while 8 percent chose one that did not even address the assignment.
- When asked to narrow an overly broad Internet search, only 35 percent of students selected the correct revision.

- When creating a presentation designed to persuade, 80 percent of study participants included both relevant and irrelevant points; 8 percent included solely irrelevant points. Just 12 percent used only those points that were directly related to their argument.

The cross-country consortium

To bridge this gap between know-how and literacy, ETS formed a consortium in 2003 with seven college and university systems, representing about 25 percent of U.S. college students.

The consortium itself was an outgrowth of a process begun two years earlier, when ETS convened an International ICT Literacy Panel to study the growing importance of information and communication technologies and their relationship to literacy. The members agreed that little was being done to address critical ICT literacy skills in higher education.

A group of ICT literacy experts assembled to advise ETS test developers on an Internet-delivered assessment to measure students' abilities to research, organize and communicate information using technology.

High-tech technology test

The group developed tasks that would identify whether students possess this knowledge. The process followed ETS's Evidence-Centered Design methodology, in which test designers first determine what they want to learn about test takers from an exam, and then create the test to precisely target that skill or knowledge.

The result of this joint effort was the *ICT Literacy Assessment*, which ETS first offered

to U.S. colleges and universities in 2005. The *ICT Literacy Assessment* presents test takers with scenario-based tasks that model the integration of content and technology that occurs — or that should be occurring — in classrooms. Tasks include selecting text and graphics that support a point of view, extracting information from a database, and composing an e-mail based on findings.

The exam not only tests for technological literacy, it's technologically advanced itself. "This is a particularly exciting assessment because it meets a critical need in education while also exhibiting the very standard for which it is testing," says ETS's Senior Vice President for Higher Education, Mari Pearlman.



"Items in the *ICT Literacy Assessment* are scenario based, not multiple-choice based, and test takers see the results of their choices in real time," Pearlman says. "So by using technology in a content-based environment, the test truly enables students to immerse themselves in its tasks."

Senior Research Scientist Katz says students' initial reactions were enthusiastic, too: Students said the test was actually enjoyable to take and that it accurately reflected real-life scenarios.

But feedback also indicated that the two-hour testing time was too long and that community college students found the assessment too difficult. ETS revised the test accordingly and re-released it in time for the 2005-06 school year.

Now 75 minutes long, the *ICT Literacy Assessment* is available in two versions: a core test that helps determine if high school seniors and college freshmen are ready for the rigors of postsecondary work, and an advanced test for community college transfers and college juniors who are ready for the higher-level demands of ICT literacy coursework.

Educators use data generated by the test to plan curricula, influence resource allocation and provide evidence for accreditation.

An information reservoir

To determine a test taker's ICT literacy, the test assesses proficiency in several areas. It evaluates students' ability to define an

information need and to access, evaluate, manage, integrate, create and communicate information.

School administrators and faculty receive score reports that provide data on individual students. Students receive individualized reports that let them know

how their overall score compares with other students' scores and how they performed on the various tasks.

"The test is as good as, if not better than, I dreamed such an assessment would be," says Lorie Roth, Assistant Vice Chancellor for Academic Programs at California State University. CalState has been studying ICT literacy since 1995 and administered the *ICT Literacy Assessment* across its 23 campuses when ETS first released it.

"When the test launched in 2005, I felt it was important enough to require all 23 campuses to take it," Roth says. "This year, we had lots of volunteers. That's a real testament to the assessment — that when people become familiar with it, they don't care if it is being required. For me, it is a dream come true."

Egan, ETS's ICT Project Manager, agrees the test is meeting a need: 60 colleges across the United States have already administered it to 10,000 students. She is also seeing a change in the educational landscape, with many schools starting to incorporate ICT literacy skills into all classes.

"This trend is similar to the integration of English and writing into other classes," Egan says. "In science classes, for example, teachers emphasize the importance of a well-written study. I think we're seeing the same phenomenon with ICT literacy."

Expanding the test's reach

The original consortium of institutions has also evolved into a broader National Advisory Committee of the following U.S. colleges and universities:

- California State University System
- City University of New York
- New Jersey Institute of Technology
- Oakland Community College, Michigan
- Purdue University
- Tarrant County College District, Texas
- University of California, Los Angeles
- University of Central Florida
- University of Memphis
- University of Texas
- University of Wisconsin

The committee advises ETS on future iterations of the *ICT Literacy Assessment* and on next steps.

International interest in the ICT assessment is also high, with ETS exploring possible uses by the University of Toronto and the United Arab Emirates. And at an Information Literacy Summit in Washington, D.C., in October 2006, for which ETS served as a co-partner, Abdelaziz Abid, UNESCO's point person on information literacy, provided a global perspective on the importance of ICT literacy.

"In some ways, you could say the U.S. is behind," says Patricia Breivik, Chair Emerita of the National Forum on Information

Literacy, a co-sponsor of the October literacy summit. Educators elsewhere have been studying the broader topic of information literacy for a decade, Breivik says.

"But ETS is really in the forefront in this country regarding companies that are paying attention to the full range of needs, instead of concentrating on the technology," she says. "The technology keeps changing, but basic critical-thinking and analytical skills will always be important."

Consistent with the literacy summit's focus on global competitiveness, ETS is working to adapt the *ICT Literacy Assessment* for use in the workplace. "The higher education assessment received a lot of media coverage, and employers recognized that the skills covered on the test were indeed the ones they needed in employees," says Egan, who is leading the workplace-adaptation project.

In interviews and surveys, employers say they want better ways to identify training needs and craft training programs. Egan says most would like to have a version of the ICT test with tasks that are content-specific to their field, and would prefer scores to be competency based rather than scaled, or averaged based on the degree of difficulty among tasks or items, as they currently are.

"The possibilities associated with a workplace adaptation are vast," says Egan, who expects the workplace test to become available in 2007. "For example, schools may want to use a version to identify training needs for their teachers — who in turn can help students strengthen their ICT literacy skills."

For more information about ETS's *ICT Literacy Assessment*, go to www.ets.org/ictliteracy. 🌟

The Results Are In

The first administrations of ETS's *Information and Communication Technology Literacy Assessment* drew praise from test takers and provided a glimpse of students' ability to use digital technology, communication tools, and networks to solve information-oriented problems.

More than 6,400 students and 65 institutions participated in the first assessments, which were administered during the first and second quarters of 2006. Among the 1,400 students who piloted the test:

- 🌟 90% said they had never taken a test like it.
- 🌟 89% described it as "appropriately challenging."
- 🌟 94% said the test required both critical thinking and technical skills.

"This is important feedback," says Teresa Egan, ETS's ICT Literacy Project Manager. "It shows that the test takers recognize the value of the test and its ability to measure important capabilities."

In addition to providing administrators and faculty with an understanding of the cognitive and technical proficiencies of students, the assessments also provided a first glimpse at the technological literacy of a generation of students raised on computer games and laptops.

"As we continue to administer the assessment, we will be able to get a clearer picture of ICT literacy in general as our score users focus on individual student scores and their implications for institutional programs and individual test takers," says Egan. 🌟

The new test measures the four language skills that are important for effective communication — reading, writing, listening and speaking — and it does so in an integrated manner so that we can determine a student's language skill level on the basis of tasks that must be performed in classrooms and on campus every day.

Benefits for score users

A test like the new TOEFL iBT provides score users a better, fairer, more complete view of a test taker's English-language proficiency than has been available ever before. To do well on it requires test takers to learn English the way it's actually used, a reality that will lead to important and fundamental changes in the way English is taught.

The new TOEFL iBT advances the important role of the test as a learning tool by providing diagnostic feedback on test performance — information that's important to test takers and teachers and that can guide and facilitate instruction.

If we can provide integrated measurements of English-language skills that mimic the way language is used in classrooms and on campuses, there is no reason to believe that we couldn't test a prospective lawyer's skills, for example,

by measuring that candidate's ability to read case law, write a brief, ask questions, and then present an oral argument.

The Internet enables us to offer significant improvements in test scoring and reporting services. We can provide students with scores and performance feedback more quickly than ever before. We can create score reports that not only inform students' academic choices but also help them understand their strengths and weaknesses so they can improve.

In the 21st century, the best assessments will not only be judged by whether they provide reliable data, but also by whether they are part of a learning system that supports teachers in the best pedagogical methods and encourages students to learn by offering them an authentic learning assessment worth having instructors teach to and students learn from.

Stand-alone assessments will almost certainly always be with us, but I believe the paradigmatic assessments of the future will be part of a system that supports learning, indeed, as the new TOEFL test does.

Research Resource

Search the ETS Research Report Series Database Online. The ReSEARCHER database contains abstracts for all the research reports in the ETS Research Report Series, which dates to 1948 and contains more than 3,000 entries. Many of the newer research reports are also available in PDF. The database is searchable by author, title, ETS research report number, publication year, key word, and text search.

<http://search.ets.org/custres/>

Professional Opportunities

ETS operates a variety of internship, fellowship and visiting scholar programs appropriate for all career levels. One aim of these programs is to increase the number of women and underrepresented minority professionals conducting research in educational measurement and related fields. Registration for most programs opens in late October; others accept applications year-round. For more information, write to

collegerelations@ets.org

ETS Classifieds

Reports and Publications

15



High School Reform and Work: Facing Labor Market Realities (Paul E. Barton, 2006)

— This ETS Policy Information

Report examines information on the abilities and qualities needed to enter the workforce with a high school diploma, including the kinds of skills and abilities that employers say they want.

www.ets.org/hsreformandwork.html



ETS Policy Notes: Latino Achievement in the Sciences, Technology, Engineering, and Mathematics (2006)

— This ETS Policy Notes provides

highlights of a conference sponsored by ETS, the College Board and the Hispanic Research Center at Arizona State University. The conference focused on promoting the success of Latino students in the STEM disciplines.

www.ets.org/latinoachievement.html



Automated Scoring of Complex Tasks in Computer Based Testing, (David Williamson, Robert Mislevy and Isaac Bejar, Eds., 2006)

— This volume reviews the latest methods of automated scoring for complex assessments and provides examples of best practices in the design, implementation and evaluation of automated complex assessment. Published by Lawrence Erlbaum Associates, the book is available for online ordering from leading book retailers.



Automated Essay Scoring: A Cross-Disciplinary Perspective (Mark Shermis and Jill Burstein, Eds., 2003)

— This book offers readers an inside look at the various automated essay scoring technologies that existed at the time of publication. Published by Lawrence Erlbaum Associates, the book also addresses pedagogical and measurement issues. The book is available for online ordering from leading book retailers.



ETS Innovations brings you news,
insight and information on educational
assessment in the United States and
around the world, from research and
test design, administration, scoring and
reporting, to test use in and out
of the classroom.



Rosedale Road
MS 30-N
Princeton, NJ 08541

STANDARD MAIL
US POSTAGE
PAID
EDUCATIONAL
TESTING SERVICE